

# Comment on: Inconsistency of the nonstandard definition of work

**Luca Peliti**

Dipartimento di Scienze Fisiche, Università “Federico II”  
Complesso Monte S. Angelo, 80126 Napoli (Italy)

**Abstract.** The objections raised by Vilar and Rubi [cond-mat arXiv:0707.3802v1] against the definition of the thermodynamical work appearing in Jarzynski’s equality are shown to be misleading and inconsistent.

Quid dices de primariis huius Gimnasii philosophis,  
qui aspidis pertinacia repleti, licet me ultro dedita  
opera millies offerente, nec Planetas, nec ☽, nec  
perspicillum videre voluerunt? Verum ut ille aures,  
sic isti oculos, contra veritatis lucem obturarunt.

*Galileo (in a letter to Kepler, 1610)*

## 1. Introduction

In a recent post, Vilar and Rubi (VR) [1] ascribe to Imparato and Peliti [2] the claim that the standard definition of work

$$\text{Work} = \text{Force} \times \text{Displacement}, \quad (1)$$

should be unconditionally replaced by a ‘nonstandard’ definition

$$dW_{\text{IP}} = -x df, \quad (2)$$

in which the force and the displacement have their role interchanged, when considering the work performed by a force on a system. They argue that “this ‘nonstandard’ definition of work is thermodynamically inconsistent at both the microscopic and macroscopic scales and leads to non-physical results, including free energy changes that depend on arbitrary parameters”. The dispute arose from the claim set forth by Vilar and Rubi in a previous post [3], in which it was argued that the connection between the microscopic work  $W$  performed by a time-dependent force on a system cannot be used to estimate free energy changes.

In the present note I shall argue the following points, which are already clear to any honest reader of ref. [2]:

- (i) The expression (2), surprising as it is, is a straightforward consequence of the standard definition of the *thermodynamical* work performed on a system, for the special case considered in [2], namely, when a uniform but time-varying force is applied to a particle subject to a given potential;

- (ii) This expression yields, via a straightforward application of the First Principle of thermodynamics, a correct evaluation of the free-energy change of a thermodynamical system undergoing a reversible transformation;
- (iii) The ‘inconsistencies’ claimed by VR to be produced by this expression of the work correspond to *bona fide* energy differences which have observable consequences.

I shall also argue that VR’s confusions stem from the fact that the thermodynamical work on a system represents the work done by the system one considers on the external bodies which act on it, rather than the work done on the system itself by the external bodies: a point stressed, *e.g.*, at the beginning of Gibbs’s founding book [4] on Statistical Mechanics, and that VR fail to appreciate.

I shall first discuss these points in the context of equilibrium thermodynamics. Further points are relevant when considering manipulated systems, in particular small systems for which fluctuations are important.

## 2. Reversible work on the harmonic oscillator

Let us consider a simple thermodynamical system, *i.e.*, a one-dimensional oscillator characterized by its mass  $m$  and spring constant  $k$ , kept at a fixed temperature  $T$ . The system is described by the hamiltonian

$$H(p, x) = \frac{p^2}{2m} + \frac{1}{2}kx^2. \quad (3)$$

In the following we shall focus only on the *displacement* degree of freedom, namely  $x$ . Its equilibrium distribution is given by

$$p^{\text{eq}}(x) = \frac{e^{-kx^2/2k_B T}}{Z}, \quad (4)$$

where  $Z$  is given by

$$Z = \int dx e^{-kx^2/2k_B T} = \sqrt{2\pi k_B T/k}. \quad (5)$$

We shall now apply a uniform, but time-varying, force  $f(t)$  to the system. We wish to evaluate the thermodynamical work performed on it, as the applied force changes from  $f_0 = 0$  to  $f$ , so slowly, that *the system can be considered to remain at thermodynamical equilibrium at all times*. This is called the *reversible work* in thermodynamics.

Following the method described by J. W. Gibbs [4] and R. C. Tolman [5], one proceeds as follows:

1. One writes down the hamiltonian of the system in the presence of the applied force:

$$H(x, f) = \frac{1}{2}kx^2 - f(x - \gamma). \quad (6)$$

Here  $\gamma$  is defined as the point in which the potential of the applied force vanishes. This point might depend on  $f$ , but we shall momentarily assume that it is fixed. It is determined by the actual device used to apply the constant force on the system, as discussed in the following.

2. One applies either Gibbs's equation (117) [4, p.45], or Tolman's equation (124.1) [5, p.542], to obtain the thermodynamical work  $dW$  associated with a small variation  $df$  of the applied force:

$$dW = \left\langle \frac{\partial H}{\partial f} \right\rangle df = -\langle (x - \gamma) \rangle df = -(\langle x \rangle - \gamma) df. \quad (7)$$

In this equation,  $\langle A \rangle$  is the canonical average of the function  $A(x)$ :

$$\langle A \rangle = \frac{1}{Z} \int dx A(x) e^{-H(x,f)/k_B T}. \quad (8)$$

In our case, one obtains

$$\langle x \rangle = \frac{f}{k}, \quad (9)$$

from which  $dW$  can be calculated via equation (7).

3. One integrates the result with a variable force  $f'$  from the initial value  $f_0 = 0$  to the final value  $f$ , obtaining

$$\Delta F = \int_0^f \left\langle \frac{\partial H}{\partial f} \right\rangle_{f'} df' = - \int_0^f \left( \frac{f'}{k} - \gamma \right) df' = -\frac{f^2}{2k} + \gamma f. \quad (10)$$

In this expression,  $\Delta F$  is the change in the Helmholtz free energy,  $F = E - TS$ . Since it is easy to see that in the present system the entropy  $S$  does not change during the manipulation, we can equate it with the change in the *internal* energy  $E$ . We have therefore

$$\Delta E = -\frac{f^2}{2k} + \gamma f. \quad (11)$$

4. Since the average value of the applied potential is given by

$$\langle U \rangle = -f(\langle x \rangle - \gamma),$$

by subtracting it from the above result, we obtain the change of the energy of self-interaction of the spring

$$\Delta E^{\text{int}}(f) = \Delta E - \langle U \rangle = \frac{f^2}{2k}. \quad (12)$$

It would not be necessary to consider this elementary exercise in statistical mechanics, were it not for the fact that in their recent post J. Vilar and M. Rubi [1] (objecting to a similar derivation contained in [2]) have found that this result is “inconsistent and unphysical both at the macroscopic and microscopic level.” Equation (7) is the one that VR incriminate. The two authors are chagrined by the following facts:

1. Let us first consider  $\gamma = 0$ . Then the free-energy change (10) is negative. Now, non-spontaneous processes should lead to positive free-energy changes. This is in contrast with previous results, including ones on macromolecules [6]. Moreover this result holds for any system described by the hamiltonian (6), including a macroscopic spring. This is in contrast with the results of elementary physics.
2. Moreover, VR claim that the parameter  $\gamma$  does not have any physical interpretation, and that therefore in this result the free-energy change does not depend on the actual physical system but rather on its mathematical description.

VR notice that the free-energy change given by equation (10) is negative since it also contains the potential energy associated with the external force. They claim that it is inconvenient to have the particular properties of the applied external force embedded into the results. Therefore they deviate from the definition of the thermodynamical work given by Gibbs, who explicitly states [4, p.4, footnote] that the energy function of the statistical system should include “that energy which might be described as mutual to that system and external bodies”. It lies on them, therefore, the burden to show that *their* ‘nonstandard’ definition of the thermodynamical work is preferable to Gibbs’s and Tolman’s one. They shun this burden by failing to notice it.

They should however agree that, if the potential energy of the interaction of the system with the bodies that provide the constant force is taken into account, both objections raised above disappear. VR proceed instead as if the expression (11) contained only the energy of interaction of the system with itself.

I now show how the result (11) corresponds to the variation in the total energy (as defined in the above text by Gibbs) when the external force is applied by two physically reasonable devices. I shall then discuss why the apparent paradox of point 1. is such only in the minds of the authors of ref. [1] and their followers. But I now wish to stress the point which probably lies at the heart of VR’s confusion, by quoting at length from Gibbs’s treatise.

Returning to the case of the canonical distribution, we shall find other analogies with thermodynamics systems, if we suppose, as in the preceding chapters,† that the potential energy ( $\epsilon_q$ ) depends not only upon the coordinates  $q_1 \dots q_n$  which determine the configuration of the system, but also upon certain cöordinates  $a_1, a_2$ , etc. of bodies which we call *external*, meaning by this simply that they are not to be regarded as forming any part of the system, although their positions affect the forces which act on the system. The forces exerted by the system *on these bodies*‡ will be represented by  $-d\epsilon_q/da_1, -d\epsilon_q/da_2$ , etc., while  $-d\epsilon_q/dq_1 \dots -d\epsilon_q/dq_n$  represent all the forces acting upon the bodies of the system, including those which depend upon the position of the external bodies, as well as those which depend only upon the configuration of the system itself. It will be understood that  $\epsilon_p$  depends only upon  $q_1, \dots q_n, p_1, \dots p_n$ , in other words, that the kinetic energy of the bodies which we call external forms no part of the kinetic energy of the system. It follows that we may write

$$\frac{d\epsilon}{da_1} = \frac{d\epsilon_q}{da_1} = -A_1, \quad (104)$$

although a similar equation would not hold for differentiation relative to the internal cöordinates.

Thus Gibbs’s expression of the elementary reversible work

$$dW = - \sum_i \bar{A}_i da_i, \quad (13)$$

(where, in Gibbs’s notation, the bar denotes the average over a canonical distribution) represents the average work *done on the external bodies by the system* (with changed sign), and therefore, in particular, does not vanish even if the coordinates of the system do not change over the time interval considered.

† See especially Chapter I, p. 4 (Note by JWG).

‡ My italics (LP).

*Electrostatic device*

To illustrate this point, let us set up a device for applying a uniform but time-dependent force on our oscillator. We can use, for instance, the following electrostatic device. Let us assume that the mass of the oscillator carries a small charge  $q$ . We take two point-like bodies at infinity, one with the charge  $+Q$  and the other with the charge  $-Q$ . We then let these two charged bodies come closer and closer to the origin (the equilibrium point of the oscillator), by letting the charge  $+Q$  be situated at the point  $-X + \gamma$ , and the charge  $-Q$  at the point  $X + \gamma$ . Thus the electric field acting on the oscillator at point  $x$  is given by

$$\begin{aligned} E &= \frac{Q}{4\pi\epsilon_0} \left[ \frac{1}{(x - X - \gamma)^2} + \frac{1}{(x + X - \gamma)^2} \right] \\ &= \frac{Q}{2\pi\epsilon_0} \frac{(x - \gamma)^2 + X^2}{[(x - \gamma)^2 + X^2]^2 - 4(x - \gamma)^2 X^2} \\ &\simeq \frac{Q}{2\pi\epsilon_0} \left\{ \frac{1}{X^2} + \frac{3(x - \gamma)^2}{X^4} + \frac{5(x - \gamma)^4}{X^6} + \dots \right\} \end{aligned} \quad (14)$$

If  $X$  is large enough, then all terms beyond the first one are negligible, for the expected excursions of the oscillator from the origin. Then the force applied by the charge  $Q$  is given by

$$f = \frac{qQ}{2\pi\epsilon_0} \frac{1}{X^2}. \quad (15)$$

Let us choose  $Q$  such that, even for the largest force  $f_1$  which we wish to apply,  $X$  is so large that the terms beyond the first in equation (14) are negligible. Thus by moving the charges  $\pm Q$  from infinity to  $\pm X + \gamma$ , always symmetrically around the point  $\gamma$ , we can apply a uniform but time-varying force to our oscillator. It is now clear that  $\gamma$ , far from being a fictitious parameter, corresponds to the location of the center of the device by which a uniform force is applied to the system we are studying. In order to change  $\gamma$ , external work must be supplied to the apparatus.

Let us now evaluate the internal energy of the system as a function of  $X$ . We have

$$\begin{aligned} E &= \left\langle \frac{1}{2} kx^2 + U(x, X) \right\rangle \\ &= \frac{1}{Z} \int dx e^{-H(x, X)/k_B T} \left[ \frac{1}{2} kx^2 + \frac{qQ}{4\pi\epsilon_0} \left( \frac{1}{x + X - \gamma} - \frac{1}{X - x + \gamma} \right) \right]. \end{aligned} \quad (16)$$

The first term yields

$$\left\langle \frac{1}{2} kx^2 \right\rangle = \frac{1}{2} k \left[ \langle (x - \langle x \rangle)^2 \rangle + \langle x \rangle^2 \right] = \frac{1}{2} k_B T + \frac{1}{2} \frac{f^2}{k}. \quad (17)$$

The first term is given by the equipartition theorem, and the second by equation (9). One can expand the second term in powers of  $1/X$ , obtaining

$$\langle U(x, X) \rangle = -\frac{qQ}{2\pi\epsilon_0} \left[ \frac{1}{X^2} \langle (x - \gamma) \rangle + \frac{1}{X^4} \langle (x - \gamma)^3 \rangle + \dots \right]. \quad (18)$$

Thus, if  $\langle (x - \gamma)^2 \rangle / X^2 \ll 1$ , we have

$$\langle U(x, X) \rangle = -\frac{qQ}{2\pi\epsilon_0} \frac{\langle (x - \gamma) \rangle}{X^2} = -\frac{f^2}{k} + \gamma f, \quad (19)$$

where we have exploited (15) and (9). Summing up, we obtain

$$E = \frac{1}{2}k_B T - \frac{f^2}{2k} + \gamma f, \quad (20)$$

in agreement with equation (11).

### Gravity-field device

A simpler conceptual experiment can be set up imagining that the oscillator mass is constrained to move along a line, which can be rotated in the vertical plane. Let  $m$  be the oscillator mass,  $g$  the acceleration of gravity, and let the hinge be placed at  $x = \gamma$ . If the line is now rotated clockwise by an angle  $\theta$ , the oscillator mass will be acted upon by a uniform force, directed towards increasing values of  $x$ , and of intensity  $mg \sin \theta$ . On the other hand, if the mass is at location  $x$ , its height with respect to the horizontal line passing through the hinge is given by  $z = -(x - \gamma) \sin \theta$ . It is then a simple matter to evaluate the average of  $U(x, \theta)$ :

$$\langle U(x, \theta) \rangle = mg \langle z \rangle = -mg \sin \theta \langle x - \gamma \rangle = -\frac{f^2}{k} + \gamma f. \quad (21)$$

Adding to it the average elastic energy  $\frac{1}{2}k \langle x^2 \rangle$  we recover equation (11) again. But it is amusing to verify that this result corresponds indeed to the work done by the system on the external device. Let us consider the line to be tilted by  $\theta$ , and the position of the oscillator to be  $x$ . Then the oscillator applies to the rectilinear guide a torque

$$\tau = mg \cos \theta (x - \gamma). \quad (22)$$

As the angle changes by  $d\theta$ , this torque executes on the guide a work

$$\tau d\theta = mg(x - \gamma) d \sin \theta.$$

The *reversible* elementary work made by the system on its environment is given by the average of this expression, namely

$$-dW^{\text{rev}} = \langle \tau \rangle d\theta = mg(\langle x \rangle - \gamma) d \sin \theta, \quad (23)$$

where, according to equation (9),  $\langle x \rangle = f/k = mg \sin \theta / k$ . The change in the internal energy due to the transformation is given by  $dW^{\text{rev}}$ , integrated between 0 and the final value of  $\theta$ . It is easy to check that it yields again the result (11).

When the rectilinear guide is tilted, the oscillator spring is stretched and its elastic energy is increased. On the other hand, the potential energy of the mass in the gravity field can either increase or decrease, and the resulting total energy change can be of either sign. If  $\gamma = 0$ , one has, for instance

$$\Delta E = -\frac{m^2 g^2 \sin^2 \theta}{2k} = -\frac{f^2}{2k}. \quad (24)$$

VR claim that this result is inconsistent, because a negative free-energy change (which coincides in our case with the energy change) would imply that the process is spontaneous, and that the spring is unstable, in contradiction with elementary physics. They fail to notice, however, that, *if the rectilinear guide is free to rotate around the origin*, the system is indeed unstable: the guide would rotate till it reaches a vertical stand, with the oscillator mass hanging on the spring. Thus, far from being unphysical, the result yields the correct prediction for the physical setup one is considering. Of course, in an actual experiment, one would *constrain* the guide at a given angle  $\theta$ , and the oscillator will find equilibrium around a point  $\langle x \rangle$  given by equation (9).

### 3. Reversible and fluctuating work

The textbook definition of reversible work is the work performed when the thermodynamic transformation is so slow that the system can be considered to stay at thermodynamic equilibrium at all times. In this case, the trajectory average coincides with the ensemble average, at least if equilibrium statistical mechanics holds. Then the performed work *does not fluctuate*, and one trivially has

$$W^{\text{rev}} = \langle W^{\text{rev}} \rangle = -k_B T \log \left\langle e^{-W^{\text{rev}}/k_B T} \right\rangle. \quad (25)$$

VR claim that this equality does not hold, presumably because in their mind the reversible work (which is a canonical average) fluctuates. On the other hand, a clear distinction was made in ref. [2] between *reversible* and *fluctuating* work, a distinction that VR chose to ignore. For the benefit of the reader, I recall the definition of the infinitesimal fluctuating work on a system whose microscopic state is denoted by  $x = (x_i)$ , and described by the hamiltonian  $H(x, \mu)$ , depending on an external parameter  $\mu$ :

$$dW = \frac{\partial H(x, \mu)}{\partial \mu} d\mu. \quad (26)$$

We then have, for a given infinitesimal change  $d\mu$ ,

$$dW^{\text{rev}} = \langle dW \rangle, \quad (27)$$

where the average is taken with respect to the canonical distribution with the given value of  $\mu$ . Notice that the fluctuating work does not depend on the *change* in the microscopic state  $x$  of the system, but on the *change of the external parameter*  $\mu$ , because it represents the work done by the system on the external bodies that act on it. One can thus understand why it does not vanish if  $\mu$  is suddenly changed: if, *e.g.*, we suddenly push the charges  $\pm Q$  closer to the origin in the electrostatic device, we have to provide some work, part of which changes the interaction energy of the oscillator with the charges. By the same token, if we change  $\gamma$ , *e.g.*, by rigidly displacing the field-creating charges  $Q$ , we have to provide work on the system, even if the oscillator's mass does not move.

The distribution of the fluctuating work exhibits a number of interesting properties, among which the remarkable equality  $\langle e^{-W/k_B T} \rangle = e^{-\Delta F/k_B T}$  derived by Jarzynski [7], and which Hummer and Szabo [8] showed how to exploit in order to obtain information on the *equilibrium* free-energy landscape from nonequilibrium experiments. VR object to this development, claiming that the above definition of the fluctuating work is unphysical and inconsistent. We have just seen how nicely it fits with equilibrium statistical mechanics, as defined by Gibbs and explained by Tolman. However, other quantities also exhibit remarkable distributions. Let us consider a system described by the hamiltonian

$$H(x, \mu) = H_0(x) - \mu F(x). \quad (28)$$

Then the fluctuating work defined above is given by

$$dW = -F(x) d\mu, \quad (29)$$

and satisfies Jarzynski's equality. On the other hand, we can also define the work  $dW_0$  by

$$dW_0 = \mu \sum_i \frac{\partial F}{\partial x_i} dx_i, \quad (30)$$

which represents the work done by the environment on the system. As recently discussed by Jarzynski [9] in more detail, this work satisfies an identity found long ago by Bochkov and Kuzovlev [10], namely

$$\left\langle e^{-W_0/k_B T} \right\rangle = 1. \quad (31)$$

However, it is *true* that it is difficult to exploit this identity in order to recover information on an equilibrium quantity like  $\Delta F$ . Indeed, what VR have brilliantly shown in [3] is that  $W_0$  cannot be used to reconstruct free-energy landscape, but they fail to inform the reader of [1] that their arguments concern the use of  $W_0$ , leaving the impression that their objections concern  $W$  and the use of the Jarzynski equality. Now there is no problem in applying the Jarzynski equality to  $W$ . The resulting  $\Delta F$  contains a contribution from the interaction between the system and the environment which, contrary to VR's statements, is easily subtracted off (see, e.g., the "histogram method" discussed in [11, 12]). It is the responsibility of the researcher to choose the most appropriate tools for one's task. One should choose a spoon to eat one's soup and a spade to dig a hole: VR appear to prescribe everybody to pick up the spoon and then they lament that it is not possible to dig holes.

#### 4. Conclusions

We have seen that VR's objections against ref. [2] stem from a biased and misleading reading of it, and from their failure to appreciate some basic concepts in statistical mechanics. I am at a loss to understand why as serious and competent physicists as VR could fall in such blunders, unless their confusions arise from an aprioristic hostility to the recent exciting developments in the statistical mechanics of manipulated systems. In this case, they would remind of Galileo's colleagues, cited in the letter I have posted *in limine*, who refused to look in the telescope because it did not fit within their world view. If it is so, let them be happy to encourage their followers to raise objections based on even faultier arguments than their own [13]. I shall have no more to say on their subject.

#### Acknowledgments

I thank A. Imparato for many discussions and insights on this matter.

#### References

- [1] J. M. G. Vilar and J. M. Rubi, cond-mat arXiv:0707.3802v1 (2007).
- [2] A. Imparato and L. Peliti, cond-mat arXiv:0706.1134v1 (2007).
- [3] J. M. G. Vilar and J. M. Rubi, cond-mat arXiv:0704.0761v2 (2007).
- [4] J. W. Gibbs, *Elementary Principles in Statistical Mechanics* (New Haven: Yale U. P., 1902), reprinted (Woodbridge CT: Ox Bow Press, 1981).
- [5] R. C. Tolman, *The Principles of Statistical Mechanics* (Oxford: Oxford U. P., 1938), reprinted (New York: Dover, 1979).
- [6] J. Liphardt et al., *Science* **296**, 1832 (2002).
- [7] C. Jarzynski, Phys. Rev. Lett. **78**, 931 (1997).
- [8] G. Hummer, A. Szabo, Proc. Natl. Acad. Sci. USA **98**, 3658 (2001).
- [9] C. Jarzynski, C. R. Physique **8**, 495 (2007).
- [10] G. N. Bochkov and Yu. E. Kuzovlev, Physica A **106**, 443 (1981).
- [11] O. Braun, A. Hanke, and U. Seifert, Phys. Rev. Lett., **93** 158105 (2004).
- [12] A. Imparato and L. Peliti, JSTAT P03005 (2006).
- [13] See, e.g., I. Santamaria-Holek and A. Perez-Madrid, cond-mat arXiv:0705.1945 (2007).